

**What Is Claimed Is:**

- 2005220-02500
1. A method for generating an index to facilitate searching through text within a document based upon an index stylesheet associated with the document, the method comprising:
- receiving the document to be indexed;
  - parsing the document to produce a parsed document;
  - retrieving instructions for creating the index for the document from the index stylesheet associated with the document; and
  - creating the index for the document by transforming the parsed document in a manner that is specified by the instructions retrieved from the index stylesheet.
2. The method of claim 1, wherein retrieving the index stylesheet involves retrieving the index stylesheet across a network from a remote address.
3. The method of claim 1, wherein the index stylesheet is appended to the document.
4. The method of claim 1, further comprising making the index available to a search engine so that the search engine is able to scan through the index.
5. The method of claim 1, wherein the index stylesheet specifies sections of the document to skip in creating the index for the document.

005220"6E4E550

1           6.       The method of claim 1, wherein the index stylesheet specifies  
2 attributes of the document that are to be included in the index.

1           7.       The method of claim 1, further comprising receiving a plurality of  
2 additional documents to be indexed and creating indexes for the plurality of  
3 additional documents using the index stylesheet.

1           8.       The method of claim 1, wherein creating the index for the  
2 document involves tokenizing the document by partitioning text within the  
3 document into individual meaning-carrying units of text.

1           9.       The method of claim 1, further comprising prior to receiving the  
2 document, downloading and parsing an index configuration file, the index  
3 configuration file specifying the index stylesheet to be used in creating the index.

1           10.      The method of claim 1, wherein receiving the document includes  
2 receiving the document from a client at an indexing server that creates the index  
3 for the client.

1           11.      A computer-readable storage medium storing instructions that  
2 when executed by a computer cause the computer to perform a method for  
3 generating an index to facilitate searching through text within a document based  
4 upon an index stylesheet associated with the document, the method comprising:  
5           receiving the document to be indexed;  
6           parsing the document to produce a parsed document;

7 retrieving instructions for creating the index for the document from the  
8 index stylesheet associated with the document; and  
9 creating the index for the document by transforming the parsed document  
10 in a manner that is specified by the instructions retrieved from the index  
11 stylesheet.

1 12. The computer-readable storage medium of claim 11, wherein  
2 retrieving the index stylesheet involves retrieving the index stylesheet across a  
3 network from a remote address.

1 13. The computer-readable storage medium of claim 11, wherein the  
2 index stylesheet is appended to the document.

1 14. The computer-readable storage medium of claim 11, wherein the  
2 method further comprises making the index available to a search engine so that  
3 the search engine is able to scan through the index.

1 15. The computer-readable storage medium of claim 11, wherein the  
2 index stylesheet specifies sections of the document to skip in creating the index  
3 for the document.

1 16. The computer-readable storage medium of claim 11, wherein the  
2 index stylesheet specifies attributes of the document that are to be included in the  
3 index.

005220"624E560

1           17.    The computer-readable storage medium of claim 11, wherein the  
2 method further comprises receiving a plurality of additional documents to be  
3 indexed and creating indexes for the plurality of additional documents using the  
4 index stylesheet.

1           18.    The computer-readable storage medium of claim 11, wherein  
2 creating the index for the document involves tokenizing the document by  
3 partitioning text within the document into individual meaning-carrying units of  
4 text.

1           19.    The computer-readable storage medium of claim 11, wherein the  
2 method further comprises, prior to receiving the document, downloading and  
3 parsing an index configuration file, the index configuration file specifying the  
4 index stylesheet to be used in creating the index.

1           20.    The computer-readable storage medium of claim 11, wherein the  
2 document is received from a client at an indexing server that creates the index for  
3 the client.

1           21.    An apparatus for generating an index to facilitate searching through  
2 text within a document based upon an index stylesheet associated with the  
3 document, the apparatus comprising:  
4           a receiving mechanism that is configured to receive the document to be  
5 indexed;  
6           a parser that is configured to parse the document to produce a parsed  
7 document;

005220-646T560

8 an stylesheet retrieving mechanism that is configured to retrieve  
9 instructions for creating the index for the document from the index stylesheet  
10 associated with the document; and  
11 an index creation mechanism that is configured to create the index for the  
12 document by transforming the parsed document in a manner that is specified by  
13 the instructions retrieved from the index stylesheet.

1 22. The apparatus of claim 21, wherein the stylesheet retrieving  
2 mechanism is configured to retrieve the index stylesheet across a network from a  
3 remote address.

1 23. The apparatus of claim 21, wherein the index stylesheet is  
2 appended to the document.

1 24. The apparatus of claim 21, further comprising an access  
2 mechanism that is configured to make the index accessible to a search engine so  
3 that the search engine is able to scan through the index.

1 25. The apparatus of claim 21, wherein the index stylesheet specifies  
2 sections of the document to skip in creating the index for the document.

1 26. The apparatus of claim 21, wherein the index stylesheet specifies  
2 attributes of the document that are to be included in the index.

1 27. The apparatus of claim 21, wherein the receiving mechanism is  
2 configured to receive a plurality of additional documents to be indexed, and

005220" 6E4ET560

3 wherein the index creation mechanism is configured to creates index for the  
4 plurality of additional documents using the index stylesheet.

1 28. The apparatus of claim 21, wherein the index creation mechanism  
2 is configured to create the index for the document by partitioning text within the  
3 document into tokens containing individual meaning-carrying units of text.

1 29. The apparatus of claim 21, further comprising an index  
2 configuration mechanism that is configured to download and parse an index  
3 configuration file, the index configuration file specifying the index stylesheet to  
4 be used in creating the index.

1 30. The apparatus of claim 21, wherein the apparatus is contained with  
2 an indexing server that creates the index for a client.